4/PRTS

09/719757
PCT/BE99/00076

JC01 Rec'd PCT/PTO 15 DEC 2000

METHOD AND DEVICE FOR CORRECTING PROXIMITY EFFECTS

The present invention relates to a method and device for determining the exposure dose of an electron beam required per pattern position to obtain a desired pattern in a coating.

In the manufacture of the latest generations of integrated circuits use is preferably made of focused electron beams in lithographic processes instead of making use of the usual optical lithographic techniques, since these latter techniques are subject to limitations in terms of resolution as a result of diffraction of the used laser light. The resolution of the integrated circuit obtained with such electron beam lithography is greater, although it is limited by scattering of the electrons in the coating. Methods are known for minimizing scattering effects or compensating therefor in advance and thereby improving the resolution of the obtained integrated circuits.

The known methods have the drawback however that scattering effects can themselves only be minimized to a limited degree, while advance compensation according to the known method requires many calculations and therefore needs a long calculation time. For the manufacture of integrated circuits for instance a very large number of pattern points, often in the order of magnitude of 10^{10} pattern points, must be "written", while the number of calculations required for this purpose amounts to a multiple thereof. As a result a practically real time precompensation for the smearing or blurring effects cannot be implemented.

The object of the present invention is to obviate this drawback and also provide additional advantages.

The present invention therefore relates to a method for determining the precompensated pattern of exposure doses of an electron beam required per pattern

position to obtain a desired pattern in a coating on a substrate, comprising of:

- determining the smearing function (blur function) of the electron beam;

- determining the precompensated pattern with the smearing function and the desired pattern, wherein the determination is performed such that the exposure doses contain almost exclusively positive values and that the exposure doses are at least to some extent smooth relative to each other.

Since a negative value for the exposure doses of an electron beam has no physical significance and cannot therefore be realized, the determination of the exposure doses of the precompensated pattern is performed such that it assumes (almost) exclusively positive values. A smooth solution is furthermore obtained since strong oscillations in the smearing function have no physical basis but are caused solely by mathematical instability of the calculations.

In a preferred embodiment of the invention the method comprises the steps of:

a) estimating a regularization parameter;

b) determining a precompensated pattern with all pattern points of the desired pattern with the exception of a determined pattern point;

c) smearing the precompensated pattern again with the smearing function in order to predict the dose of the determined pattern point;

d) repeating steps b and c for each pattern point;

e) repeating steps a to d with adapted regularization parameter until a final value of a regularization parameter is obtained;

f) determining the precompensated pattern with the final value of the regularization parameter.

According to a further embodiment of the invention step b) comprises the following iterative definition:

with $d^{(0)} = 0$ and $r^{(0)} = a$
in which a is a vector with the doses of the desired pattern as elements, d is a vector with the exposure doses of the precompensated pattern, K is the smearing function in matrix form, K^* is the Hermitian conjugate of the smearing function K, B is an operator and λ a regularization parameter.

According to a further embodiment of the invention operator B is defined as follows:

$$B(D) = \sum_i \left(\frac{d_i}{d_{tot}}\right) \ln\left(\frac{d_i}{d_{tot}}\right)$$

in which the summation takes place over all pattern points, d_i is the i^{th} element of the vector d, and d_{tot} represents the summation over all elements of the vector d.

According to a further embodiment of the invention the final value of the regularization parameter in the above mentioned step e) is the regularization parameter wherein

$$\frac{1}{N}\sum_{k=1}^{N} (a_k - [Kd_k(\lambda)]_k)^2$$

in which N is the total number of pattern points, a is a vector with the doses of the desired pattern as elements, d is a vector with the exposure doses of the precompensated pattern and K the smearing function in matrix form.

According to a further embodiment of the invention the final value of the regularization parameter in step e) is the minimal regularization parameter wherein

$$\frac{1}{N}\sum_{k=1}^{N}(a_k - [Kd^k(\lambda)]_k)^2 w_{kk}(\lambda)$$

in which N is the total number of pattern points, a is a vector with the doses of the desired pattern as elements, d is a vector with the exposure doses of the precompensated pattern, K is the smearing function in matrix form and w_{kk} is defined as:

$$w_{kk}(\lambda) = [\frac{1-a_{kk}(\lambda)}{1-\frac{1}{N}\sum_{j=1}^{N}a_{jj}(\lambda)}]^2$$

with a_{kk} the elements of the matrix $A = K(K^TK+\lambda L(D)^TL(D))^{-1}K^T$ and L the discrete Laplace transformation.

According to a further embodiment of the invention after step e) the step is performed of training a neural network using one or more desired first patterns and the associated precompensated patterns.

According to a further embodiment of the invention the precompensated pattern associated with a second desired pattern can be determined with the trained neural network, wherein in a further embodiment the first desired pattern is a relatively simple training pattern and the second desired pattern is the partial pattern of an integrated circuit, and wherein in a further embodiment two or more partial patterns can be combined into a composite pattern of the integrated circuit.

By determining the associated precompensated pattern of exposure doses for a known desired pattern, which is preferably simple, and then establishing the relation between the weighting factors of a neural network, is ensured that for a second desired pattern, which may be complicated, obtaining the relation between this pattern and the associated exposure doses is determined in very efficient and rapid manner by the neural network. The first pattern is generally a relatively simple training pattern, while the second

pattern is for instance the pattern of a very complicated integrated circuit.

In a preferred embodiment of the invention the above stated neural network is implemented in hardware, whereby determining of the relation between a pattern and the exposure dose associated therewith is performed in more rapid manner, for instance within 60 ns for each pattern point and within 10 minutes for a pattern of 10^{10}.

According to a preferred embodiment of the invention the smearing function is made up of at least two Gaussian functions, to which an exponential function is optionally added. Parameters of the Gaussian functions and optionally the exponential function can be determined by means of statistical simulation of the system of electron beam transmitting equipment and the relevant coating and substrate of the integrated circuit for manufacture.

In another embodiment of the invention parameters are determined by performing measurements on the system of electron beam transmitting equipment and the relevant coating with substrate.

The present invention also relates to a device for determining the exposure dose of an electron beam required per pattern position to obtain a desired pattern in a coating on a substrate, comprising an electronic circuit for implementing said neural network with weighting factors determined in the above stated manner.

The invention will be elucidated hereinbelow with reference to a preferred embodiment thereof, wherein use will be made of the annexed drawings, in which:

- figure 1 shows a schematic overview of a preferred embodiment of a device according to the invention;

- figures 2a-2c show a schematic overview of the determination of a precompensated pattern of 3x3 pattern points;

- figure 3 shows a desired training pattern of 256x256 pattern points;

- figure 4 shows the training pattern of figure 3 after smearing;
- figure 5 shows a graph in which for the training pattern of figure 3 the prediction error is plotted as a function of the chosen regularization parameter;
- figure 6 shows the training pattern of figure 3 after precompensation;
- figure 7 shows the precompensated pattern of figure 6 after smearing; and
- figure 8 is a schematic representation of a neural network for determining precompensated patterns.

In an arrangement of equipment for transmitting an electron beam and a substrate 1 with coating 2 for processing, a beam of electrons 3 is directed at a position or pattern point of a coating 2 on a substrate 1. The interaction of the incident electron beam 3 with the coating or resist film 2 and the underlayer or substrate 1 results in a scattering of the electrons in coating 2 which causes smearing or proximity effects. When for instance a primary electron penetrates into the coating, a part of its energy is transferred to electrons of the atoms of the coating, which causes ionization or excitation thereof. A collision between electrons with a large transfer of energy generates a secondary electron which generally has a direction of movement perpendicular to that of a primary electron.

Smearing effects in electron beam lithography relate more generally to the process whereby the resolution of the exposed pattern is reduced by primary electron scattering (forward scattering) and secondary electron excitation (backward scattering) in the coating and the substrate of an integrated circuit for manufacture. Sharp features such as angles in the desired pattern are rounded, line thicknesses and interspaces are modified and in particular extreme cases some features even disappear completely or are merged in incorrect manner with adjacent features.

The smearing effects or proximity effects can be described by a smearing function, which shows the relation between on the one hand the exposure doses of a determined pattern point of a pattern for manufacturing in the coating and on the other the doses actually absorbed by this pattern point and adjacent pattern points. The effect of the smearing is thus established in the smearing function.

Assuming that exposure and smearing are linearly and spatially invariant and that for a numeric solution a discrete representation is preferred, the above can be expressed in matrix form as follows: A = KD, in which A is a column vector of which each element a_i is the total energy dose which is actually absorbed in the associated pattern point, K is a smearing matrix of which each mn^{th} element is the portion of the energy dose which is absorbed in pattern point m from a unit-exposure dose supplied to pattern point n, and D is a column vector made up of elements d_i which represent the exposure doses generated per pattern point by the electron beam equipment. Since the smearing effect is unavoidable, it is best to adapt the exposure doses d_i of the different pattern points such that the dose a_i actually absorbed in a pattern point is such that the desired pattern is still obtained.

This so-called precompensation of the exposure dose of the electron beam can be performed according to the prior art by determining the inverse of the smearing matrix K. There are many ways of generally inverting a matrix. However, these methods often take no account of physical limitations, such as in this case for instance those of the electron beam transmitters. No negative exposure doses for instance are thus possible. A further drawback of such inversion methods is that the inverted matrix has many oscillations. In addition, for inversion of the smearing matrix for a partial pattern of for instance 256x256 pattern points the inversion of a smearing matrix with dimensions of 65536x65536 has to be

calculated, which requires an enormous amount of calculating time.

Figures 2a to 2c show a desired pattern (A). The pattern is built up of 9 pattern points a_i wherein i varies from 1 to 9. This desired pattern must be precompensated in order to be able to provide the desired pattern after exposure to the smearing electron beam, i.e. the values of d_i, with i varying from 1 to 9, have to be determined.

The precompensated pattern is first of all determined making use of the doses a_i with i from 2 to 9, wherein pattern point 1 is not therefore taken into account (figure 2a). This precompensated pattern is determined on the basis of the following expression:

$$d^{(l)} = d^{(l-1)} + (K^{*}K+\lambda B(D))^{-1}K^{*}r^{(l-1)} \qquad r^{(l)} = a-Kd^{(l)}$$

with $d^{(0)} = 0$ and $r^{(0)} = a$

wherein a is a vector with the doses of the desired pattern as elements, d is a vector with the exposure doses of the precompensated pattern , K is the smearing function in matrix form, K^{*} is the Hermitian conjugate of smearing function K, B is an operator and λ a regularization parameter. The value of the regularization parameter can be chosen at random, in this case for instance $\lambda = 0$.

The operator B imposes a limitation and can be defined as follows:

$$B(D) = \sum_{i} \left(\frac{d_i}{d_{tot}}\right) \ln\left(\frac{d_i}{d_{tot}}\right)$$

in which the summation takes place over all pattern points, d_i is the i^{th} element of the vector d, and d_{tot} represents the summation over all elements of the vector d.

The thus determined precompensated pattern is then smeared once again on the basis of the known

smearing function, whereby the predicted dose Kd of pattern point 1 is determined.

The above procedure is then repeated successively (figures 2b and 2c) for the second to ninth pattern point (i=2,..,9), wherein all pattern points with the exception of one pattern point are used each time.

On the basis of the above results, the least squares prediction error over all pattern points is determined, which will be further explained later.

The above procedure is subsequently repeated with different values for the regularization parameter λ, for instance λ_2= 0.001, λ= 0.002 etc. The regularization parameter is eventually chosen wherein the least squares prediction error over all pattern points is minimal. This regularization parameter is then chosen as the optimal regularization parameter λ_{opt}. The final precompensated pattern is then determined on the basis of this optimal regularization parameter λ_{opt}.

For this purpose the minimum is determined of the expression:

$$\frac{1}{N}\sum_{k=1}^{N}(a_k - [Kd^k(\lambda)]_k)^2 w_{kk}(\lambda)$$

in which N is the total number of pattern points, a is a vector with the doses of the desired pattern as elements, d is a vector with the exposure doses of the precompensated pattern, K is the smearing function in matrix form and w_{kk} is defined as:

$$w_{kk}(\lambda) = \left[\frac{1-a_{kk}(\lambda)}{1-\frac{1}{N}\sum_{j=1}^{N}a_{jj}(\lambda)}\right]^2$$

with a_{kk} the elements of the matrix $A = K(K^TK+\lambda L(D)^TL(D))^{-1}K^T$ and L the Laplace operator.

The smearing function resulting from forward scattering and backward scattering of the electrons of

the electron beam can be determined in different ways. It can be determined on the basis of measurements of the impulse response of the equipment for transmitting the electron beam on a test object. The smearing function can also be determined using diverse Monte Carlo techniques. In the first method of determination all physical aspects of the equipment used are taken into account. In the latter mentioned method of determination only a model of the reality is used, although the determination is however easier to perform without requiring extensive measurements.

Gaussian functions are preferably used as approximation for the smearing functions determined in any of the above described methods. The smearing function is in this case "fitted" for instance with a scattering fit model of a double Gaussian function (for both forward and backward scattering properties of the electrons), a triple Gaussian function or a double Gaussian function with a decreasing exponential function. The choice of the scattering fit model depends on the dimensions of the components to be distinguished in the test object (resolution). At dimensions smaller than 100 nm the choice hereof becomes critical: at such small dimensions the triple Gaussian functions or double Gaussian functions with decreasing exponential function are recommended. A smearing function with double Gaussian function can be described with 3 parameters, while the other two stated scattering fit models can be described with 4 parameters, which implies a great reduction in the quantity of data for processing.

Figure 3 shows a desired pattern of 256x256 pattern points. Smearing with a smearing function in the form of a double Gaussian function with α= 50 nm, ß= 3.45 and ^{n}= 1.36 produces the smeared pattern of figure 4. It is clearly visible that much detail in the pattern has been lost, which means a limitation in the resolution to be obtained of the pattern for manufacture. Application of the method according to the invention produces an

optimal regularization parameter of λ_{opt}= 0.07042, which is shown in figure 5, in which the error in the pattern is minimal at this value of λ. The precompensated pattern calculated with this value of λ and the associated smeared pattern are shown respectively in figures 6 and 7. Comparison of the results of figure 7 with those of figure 3 shows that the precompensation of the pattern with a desired pattern produces a smeared pattern with a greatly improved resolution. Components for distinguishing with dimensions of less than 100 nm, for instance in an integrated circuit, can hereby be realized. A comparison of the results of the method described herein with those of other correction methods is shown in table 1. The degree of error of the correction methods is defined here as the summation of the difference between the calculated exposure doses and the ideal precompensated exposure doses divided by the number of pattern points.

Correction method	degree of error in %
Uncorrected	10.2 %
Truncating	10.2 %
Shifting and scaling	12.2 %
Present method	4.9 %

From the above can be seen that the present method of determining a precompensated pattern produces by far the smallest degree of error compared to the other usual methods.

The precompensated pattern and the desired pattern are subsequently used as training set or training patterns for a neural network. A part of such a network is shown schematically in figure 8 en is represented by the expression

$$a_i = \sum_{j=1}^{9} w_{ij} h_{ij}(x)$$

i.e. the dose a_i is expressed in a set of 9 basic functions h_{ij}, in this case radial functions.

After training of the neural network a precompensated pattern can be determined for another random desired pattern in very rapid manner. A random pattern can for instance be a pattern of 512 by 512 pattern points forming a partial pattern of an integrated circuit. Various partial patterns can then be combined (clustered) to form one pattern which comprises the whole integrated circuit or at least a part thereof.

The above described neural network can be implemented in hardware, and preferably in analog hardware since the calculating speed of neural networks implemented in this manner is very great. The calculating time for precompensation of a pattern thus amounts to less than 60 ns per pattern point. Precompensation of a pattern of an integrated circuit of about 10^{10} pattern points requires in this case only about 10 minutes on present personal computers.

The invention is further described in the non-prepublished doctoral thesis with the title "Proximity effects correction in electron beam nanolithography", the entire content of which should be deemed as interpolated herein.